AGU Dec. 2013 H11E-1200

Enhancing Access to and Use of NASA Hydrological Data

NASA/Goddard EARTH SCIENCES DATA and INFORMATION SERVICES CENTER (GES DISC)

Two methods for enhancing access to and use of NASA data via CUAHSI-HIS and other hydrological community tools

Hualan Rui[1,2,5], Richard Strub[1,2], William Teng[1,2], Bruce Vollmer[1], David Mocko[1,3], David Maidment[4], Timothy Whiteaker[4]

[1]NASA Goddard Space Flight Center, [2]ADNET Systems, Inc., [3]SAIC, [4]University of Texas-Austin

[5]*Email: Hualan.Rui@nasa.gov*

Introduction

The way NASA earth sciences data are typically archived (by time steps, one step per file, often containing multiple variables) is not optimal for their access by the hydrological community, particularly if the data volume and/or number of data files are large. To enhance the access to and use of these data, the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) adopted two approaches, in a project supported by the NASA ACCESS Program.

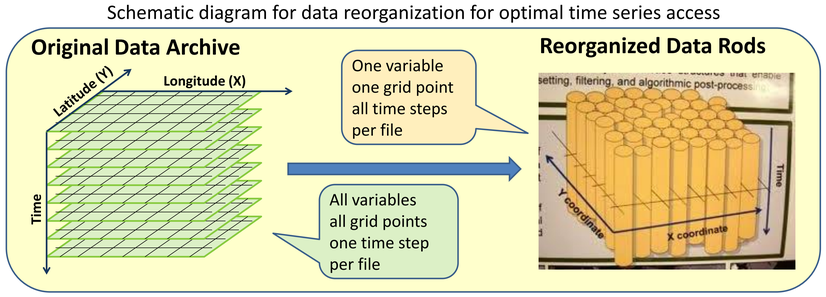
- Optimally reorganize large hydrological data sets for more efficient access, as time series (aka "data rods").
- Leverage the NASA Simple Subset Wizard (SSW) to generate parameter- and spatially- subsetted time series data.

Thus far, selected variables from the North American and Global Land Data Assimilation Systems (NLDAS and GLDAS, respectively) and the Modern-Era Retrospective Analysis for Research and Applications (MERRA) have been integrated into hydrological community tools, such as CUAHSI-HIS, EPA-BASINS, and Esri-ArcGIS.

Enhancing Access to Big Data at NASA

Approach 1: Optimal reorganization – pre-subsetted Time Series

Schematic diagram for data reorganization for optimal time series access



Advantages:

- Data rods can be retrieved, for any grid "point," any time period up to entire temporal range, with great performance.
- Data rods can be accessed as ASCII and time series plot.
- Mostly a one-time effort that benefits all users.

Disadvantages:

- Significant resources (data processing software development, disk space, CPU).
- Limited number of variables.
- Additional efforts when source data are reprocessed (corrections or new versions).

Approach 2: Leveraging SSW – Subset to Time Series On-The-Fly

NASA Simple Subset Wizard (SSW) was developed to unite data search and subsetters at various NASA EOSDIS data centers into a single, simple, seamless process. Data accessed via SSW are converted to time series before being made available via Web service.

Advantages:

- All SSW-accessible variables potentially available to HIS users.
- Much less up-front efforts and resources.
- No additional effort needed if data are reprocessed.

Disadvantages:

- For some high resolution data sets, performance is an issue.
- Length of time series accessed is limited.
- Subsetting service repeated for each request, consuming system resources.

Acknowledgment: This work is supported by NASA ROSES NNH11ZDA001N-ACCESS.

NASA Hydrological data in CUAHSI-HIS

NASA Hydrological Variables in HIS

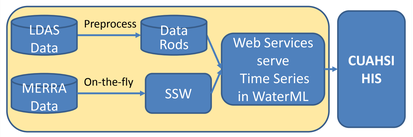
		Variable Name
NLDAS-2 Hourly 0.125°	Primary Forcing	Precipitation hourly total
		2-m above ground temperature
		10-m above ground zonal wind speed
		10-m above ground meridional wind speed
		Potential evaporation
		2-m above ground specific humidity
	Noah	0-100 cm top 1 meter soil moisture content
		0-10 cm soil temperature
		Surface runoff (non-infiltrating)
		Total evapotranspiration
GLDAS-1 3-hourly 0.25°	Noah Model	Precipitation rate
		Rainfall rate
		Snowfall rate
		0-100 cm top 1 meter soil moisture content
		0-10 cm layer 1 soil moisture content
		10-40 cm layer 2 soil moisture content
		40-100 cm layer 3 soil moisture content
		Total evapotranspiration
		Near surface air temperature
		Near surface specific humidity
		Surface runoff
		Near surface air temperature
		Near surface wind magnitude
MERRA Hourly 0.5°x0.6667°	Land Diagnostics	Surface total precipitation
		Top soil layer soil moisture content
		Root zone soil moisture content
		Total profile soil moisture content
		Top soil layer soil wetness
		Root zone soil wetness
		Total profile soil wetness
		Overland runoff
		Bare soil evaporation
		Evaporation from land
		Transpiration
		Mean land surface temperature (incl. snow)
		Soil temperature in (layer 1, 2, 3, 4, 5, and 6)

CUAHSI: Consortium of Universities for the Advancement of Hydrologic Science, Inc.

HIS: Hydrologic Information System

- CUAHSI-HIS is an internet-based system for sharing water data.
- Three components of HIS: Data Clients, HydroServer, and HydroCatalog.
- Currently total about 115 data services.
- Three of them for NLDAS, GLDAS, and MERRA-Land.

http://hiscentral.cuahsi.org/pub_services.aspx



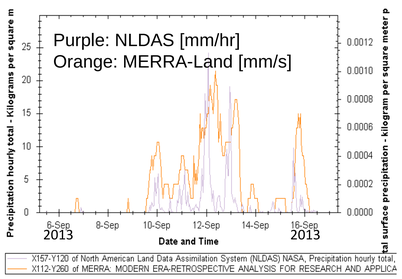
Schematic diagram for integrating LDAS and MERRA data into CUAHSI-HIS.

WaterML: Water markup language, an encoding standard for hydrologic data.

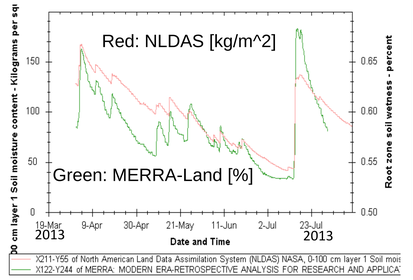
HydroDesktop is a free Windows-based GIS application for discovering, accessing, and analyzing hydrological data brokered by CUAHSI-HIS.

HydroDesktop users can access NLDAS, GLDAS, and MERRA-Land data, along with other existing HIS data, and use built-in functions for water-cycle-related research, applications, and data validation.

Sample Time Series of NLDAS and MERRA-Land inter-comparison, generated by HydroDesktop



NLDAS grid in Boulder County (105.3125W, 40.0625N)



NLDAS grid in central Texas (98.5625W, 31.9375N)

Summary

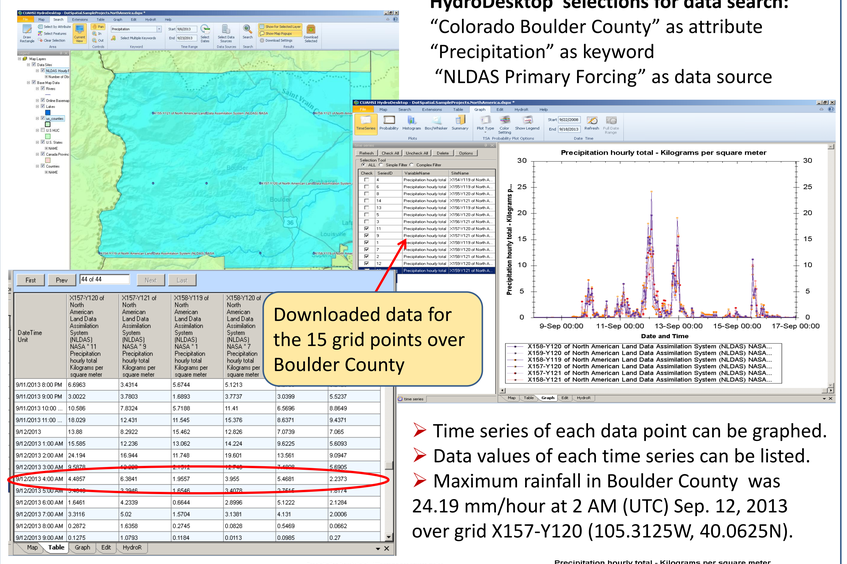
- Two methods for enhancing access to and use of NASA hydrological data via CUAHSI-HIS and other hydrological community tools are implemented by: (1) optimally reorganizing big data to data rods; (2) leveraging SSW.
- WaterOneFlow Web Services implemented at NASA GES DISC enabled integration with CUAHSI HIS by serving NASA hydrology data in WaterML.
- Selected variables from LDAS and MERRA-Land have been integrated into CUAHSI-HIS and can be accessed via HydroDesktop for water-cycle-related studies.
- Application example: Colorado flood September 2013 studied with NLDAS-2 precipitation via HydroDesktop of CUAHSI-HIS (see upper right).
- Application prototype: Data rods (GLDAS top 1 meter soil moisture) in ArcGIS Online enabled access to soil moisture data for any land location in the world (see lower right).
- TRMM precipitation (TRMM 3B42) being integrated into CUAHSI-HIS by SSW approach.

Application Examples

Colorado's "Biblical" Flood

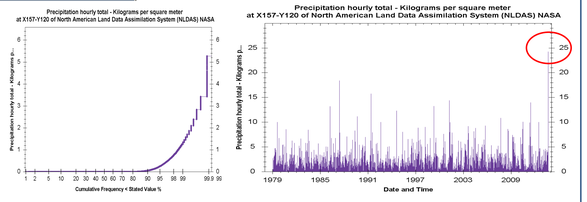
After the state experienced its worst wildfire on record earlier in 2013, Colorado was hit by record rainfall and flooding, with more than half a year's worth of rain (~8-10") falling over three days in the middle of September.

HydroDesktop selections for data search:
"Colorado Boulder County" as attribute
"Precipitation" as keyword
"NLDAS Primary Forcing" as data source



- Time series of each data point can be graphed.
- Data values of each time series can be listed.
- Maximum rainfall in Boulder County was 24.19 mm/hour at 2 AM (UTC) Sep. 12, 2013 over grid X157-Y120 (105.3125W, 40.0625N).

- Cumulative Frequency of rainfall can be calculated and graphed.
- Cumulative Frequency plot shows 99.99% of rain rate at the grid are less than 5.23 mm/hour.



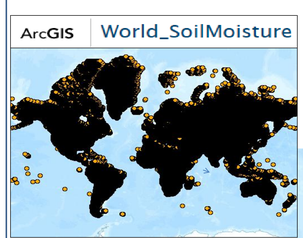
Hydrology Portal

LDAS Portal

- Time series of entire temporal range of data at the grid shows rain rate of 24.19 mm/hour is the maximum over the 33 years (1979 ~ 2013).
- NLDAS soil moisture and surface runoff can also be studied for this extreme event through CUAHSI-HIS.

World Soil Moisture Map, an example of GLDAS in ArcGIS Online

As a prototype, integration of data rods of GLDAS top 1 meter soil moisture in ArcGIS online enabled access to soil moisture data (as time series plot and ASCII) for any land location in the world. http://www.arcgis.com/home/webmap/viewer.html?webmap=7d6cefdf3f324b55b08c136654e91612



Charts show 3-hourly variation (more than 40,000 values) of soil water content of the top 1 m of soil from 2000 to 2013. Total 243,884 grid points over land.

Popup on point links to data and chart